

O'Bryen, Barbara

From: Goldberg, Jeanine
Sent: Tuesday, April 02, 2002 7:34 AM
To: O'Bryen, Barbara
Subject: 09/762,724- MSG Pneumocystis carinii

Please place results on DISK.

1. Please search SEQ ID NO: 17-20, 23, 24
2. Please search SEQ ID NO: 13, positions 2821-3072
3. Please search SEQ ID NO: 3, positions 2758-3006
4. Please search SEQ ID NO: 1, positions 2894-3042
5. Please search SEQ ID NO: 5, posigions 2845-3090

THANK YOU
Jeanine

Jeanine Enewold Goldberg
1634
CM1--12D11
Mailbox-- 12E12
306-5817

Point of Contact:
Barb O'Bryen
Technical Information Specialist
STIC CM1 6A05 308-4291

B0B
4-2-02

SEQ ID NO: 13, nucleotides 2821-3072

RESULT 5

PMCMMSGI

LOCUS PMCMMSGI 3363 bp DNA PLN 26-SEP-1994

DEFINITION Pneumocystis carinii B-cell receptor (msgI) gene, 3' end.

ACCESSION L27092

VERSION L27092.1 GI:535706

KEYWORDS B-cell receptor.

SOURCE Pneumocystis carinii DNA.

ORGANISM Pneumocystis carinii

Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; Pneumocystis.

REFERENCE 1 (bases 1 to 3363)

AUTHORS Garbe, T.R. and Stringer, J.R.

TITLE Molecular characterization of clustered variants of genes encoding
major surface antigens of human Pneumocystis carinii

JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)

MEDLINE 94314421

FEATURES Location/Qualifiers

source 1. .3363

/organism="Pneumocystis carinii"

/db_xref="taxon:4754"

mat_peptide 152. .3241

/gene="msgI"

/standard_name="major surface antigen"

/evidence=experimental

/product="B-cell receptor"

gene 152. .3244

/gene="msgI"

CDS <152. .3244

/gene="msgI"

/codon_start=1

/product="B-cell receptor"

/protein_id="AAA21645.1"

/db_xref="GI:535707"

/translation="VARAVKRQVAGVKNNAEERLFALIMRADYKDESKCKNKKIKEYC
DGLKNASLTSEEVHVELKDFCKDGKQGKCEELKKNEAKCNNFKTKLEGVLVKKDASG
LTNDDCKENERQCLFLEGACPDLVEDCSKLRNLCYQKKREGVAEEVLLRALRGDLGNK
TECEKKIKDVCVPKIGQESDETLLLCQLDKKTCTNLMTARDKKCNTLEEDVKKALENKN
NLLGKCLPLLEHATFTEGTAKKASQCTPNKDCEDYLPKCDELAEECGKKGIYIHPGP
DFDPTKPEPTVAEDIGLEELYKKAEDGVHIGKPPVRDATAALLLIQNPDPKIQANE
KEKCKKVLENCKELKKHEVLDLNCQNAASQSGTKKCELEKELANSTKILSEKIKN
KHLGSGSETIPWYKLSTFLSDSDCARLESCDFYFAQDKDPLKKECKNVKAACYKRLD
ARANKVLQENMRGLLRGSNQSWLKKFQQELVKVCEKLKEENKGSFSNDELFLCVQPA
KAARLLTHDLRMTTIFLRQQLDQKRDFPTVKTAKGIREKQD LGKFQKEITWPCHTL
EQQCNRGTTIELKQVLLNEHKDTLKTHENCVTYLKEKCNKWSRRGDRFSFVCVFQN
ATCKLMVKDVQDRCKIFKENIKVSEIVDFLNNTNNITTLERNCP SWHTYCNRFSSNC
PDFSKKNPCTKIKNNCKPFYERKALEDALKVELRGKLSDENKCTAALKGYCTLAGNVN
NASVRSLCKDNTQGSNKTDEKVV EELCKKLMEEVKEQCETLPAELKQPAD DLEKDV
TYEELKEEAKKAMNKSSLVLSFVKKDGNNTPKNNSKSEDKNVVSNEKDTIKHVKILRR
GVKDVLVTELEAKAFDLAAEVFGRYVDLKERCEKLTLDGIKDDCDGLKGVCVGKIKKK
CRDLKPLEVKSHIEVTESTTTTTTVTDPKATECKSLQTTDTWVTQTSTHTSTST
ITSTITSKITLTSTRCKPTKCTTGDDAEDVKPSEGRLVSGWNVMRGAIVAMVISFMI
"

BASE COUNT 1312 a 505 c 753 g 793 t

ORIGIN

SEQ ID NO: 3, nucleotides 2758-3006

RESULT 7

PMCMMSGI

LOCUS PMCMMSGI 3363 bp DNA PLN 26-SEP-1994
DEFINITION *Pneumocystis carinii* B-cell receptor (*msgI*) gene, 3' end.
ACCESSION L27092
VERSION L27092.1 GI:535706
KEYWORDS B-cell receptor.
SOURCE *Pneumocystis carinii* DNA.
ORGANISM *Pneumocystis carinii*
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; *Pneumocystis*.
REFERENCE 1 (bases 1 to 3363)
AUTHORS Garbe, T.R. and Stringer, J.R.
TITLE Molecular characterization of clustered variants of genes encoding
major surface antigens of human *Pneumocystis carinii*
JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)
MEDLINE 94314421
FEATURES Location/Qualifiers
source 1. .3363
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"
mat_peptide 152. .3241
/gene="msgI"
/standard_name="major surface antigen"
/evidence=experimental
/product="B-cell receptor"
gene 152. .3244
/gene="msgI"
CDS <152. .3244
/gene="msgI"
/codon_start=1
/product="B-cell receptor"
/protein_id="AAA21645.1"
/db_xref="GI:535707"
/translation="VARAVKRQVAGVKNEAEERLFALIMRADYKDESKCKNKIKEYC
DGLKNASLTSEEVHKELKDFCKDGSGKKCEELKNVEAKCNNFKTKLEGVLVKDASG
LTNDDCKENERQCLFLEGACPDLVEDCSKLRNLCYQKKREGVAEEVLLRALRGDLGNK
TECEKKIKDVCPKIGQESDETLCLLDQKKCTNLMTARDKKCNTLEEDVKKALENKN
NLLGKCLPLLEHATFTEGTAKKASQCTPNKDCEYLPCDELAEECGKKGIYIHPGP
DFDPTKPEPTVAEDIGLEELYKKAEDGVHIGKPPVRDATAALLLIQNPDPKIQANE
KEKCKKVLENKCKELKKHEVLGDLCNQNAASQSGTKKCELEKELANSTKILSEKIKN
KHLSGSGETIPWYKLSTFLSDSDCARLESDCFYFAQDKDPLKKECKNVKAACYKRGLD
ARANKVLQENMRGLLRGNSNQSWLKKFQQELVKVCEKLKEENKGSFSNDELFLCVQPA
KAARLLTHDLRMTTIFLRQQLDQKDFPTVKTAKGIREKCQDLGKGQKEITWPCHTL
EQOCNRGTTIELKQVLLNEHKDTLKTHENCVTYLKEKCNKWSRRGDRFSFVCVFQN
ATCKLMVKDVQDRCKIFKENIKVSEIVDFLKNNNTNNITTLENCPSWHTYCNRFSSNC
PDFSKKNPCTKIKNNCKPFYERKALEDALKVELRGKLSDENKCTAALKGYCTLAGNVN
NASVRSLCKDNTQGSNKTDEKVVELCKKLMEEVKEQCETLPAELKQPADDLEKDVK
TYEELKEEAKAMNKSSLVLSFVKDGNNTPKNNSKSEDKNVVSNEKDTIKHVKILRR
GVKDVLVTELEAKAFDLAAEVFGRYVDLKERCEKLTDCGIKDDCDGLKGVCVGKIKKK
CRDLKPLEVKSHEIVTESTTTTTTVTDPKATECKSLQTTDTWVTQTSTHTSTST
ITSTITSKITLTSTRRKPTKCTTGDDAEDVKPSEGRLVSGWNVMRGAIVAMVISFMI
"
" " "

BASE COUNT 1312 a 505 c 753 g 793 t
ORIGIN

Query Match 61.6%; Score 153.4; DB 8; Length 3363;
Best Local Similarity 79.1%; Pred. No. 1.2e-33;
Matches 197; Conservative 0; Mismatches 46; Indels 6; Gaps 1;

y 1 gactgccactttacagacaacagatacgtgggtcacaaagacgtcgaccatacttagc 60
|| ||| | ||||||| || || ||| || ||| ||| ||| ||| ||| ||| |||
o 2999 GAATGCAAATCCTTACAGACAACAGACACATGGTTACACAGACATCGACACACACAAGC 3058

y 61 acatccacaaccacatctacagtcacgtcaagaataacgttgcacctcgacaagacgtgt 120
||||| || | ||||||| | ||| ||| ||| ||| ||| ||| |||
b 3059 ACATCTACCATCACATCTACGATTACATCAAAAATAACATTGACATCAACAAGGCGGTGC 3118

y 121 aagcctacgaagtgtacgacaggagagaagatgaagcaggagacgtgaaaccgagtgaa 180
|| | || | ||||||| ||| ||| ||| ||| ||| ||| |||
b 3119 AAACCAACCAAGTGTACGACAGG-----GGATGATGCAGAACAGTGAAGCCAAGTGAA 3172

y 181 gggttgggatggatggatggatgttatggatgggggggtttattagcaatgacgattca 240
|| | ||| | | ||| ||| ||| ||| ||| ||| |||
b 3173 GGCTTGAGGGTGAGCGGGTGGAAATGTGATGAGGGGGCAATAGTAGCAATGGTTATTCG 3232

y 241 ttcatgatt 249
||| | | | |
b 3233 TTCATGATT 3241

SEQ ID NO: 1, nucleotides 2894-3042

RESULT 7

PMCMMSGI

LOCUS PMCMMSGI 3363 bp DNA PLN 26-SEP-1994

DEFINITION *Pneumocystis carinii* B-cell receptor (*msgI*) gene, 3' end.

ACCESSION L27092

VERSION L27092.1 GI:535706

KEYWORDS B-cell receptor.

SOURCE *Pneumocystis carinii* DNA.

ORGANISM *Pneumocystis carinii*

Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;

Pneumocystidaceae; *Pneumocystis*.

REFERENCE 1 (bases 1 to 3363)

AUTHORS Garbe, T.R. and Stringer, J.R.

TITLE Molecular characterization of clustered variants of genes encoding major surface antigens of human *Pneumocystis carinii*

JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)

MEDLINE 94314421

FEATURES Location/Qualifiers

source 1. .3363
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"

mat_peptide 152. .3241
/gene="msgI"
/standard_name="major surface antigen"
/evidence=experimental
/product="B-cell receptor"

gene 152. .3244
/gene="msgI"

CDS <152. .3244
/gene="msgI"
/codon_start=1
/product="B-cell receptor"
/protein_id="AAA21645.1"
/db_xref="GI:535707"

/translation="VARAVKRQVAGVKNEAEERLFALIMRADYKDESKCKNKKIKEYC
DGLKNASLTSEEVHKELKDFCKDGSGKKCEELKNVEAKCNNFKTKLEGVLVKKDASG
LTNDDCKENERQCLFLEGACPDILVEDCSKLRNLCYQKKREGVAEEVLLRALRGDLGNK
TECEKKIKDVCPKIGQESDETLLLCQLDKKTCTNLMTARDKKCNTLEEDVKKALENKN
NLLGKCLPLLEHATFTEGTAKKASQCTPNKDCEDYLPKCDELAEECGKKGIYYIHPGP
DFDPTKPEPTVAEDIGLEELYKKAEDGVHIGKPPVRDATAALLALLIQNPDPKIQANE
KEKCKKVLENCKELKKHEVLGDLCNQNAASQSGTKKCEELEKELANSTKILSEKIKN
KHLGSGSETIPWYKLSTFLSDSDCARLESDCFYFAQDKDPLKKECKNVKAACYKRLD
ARANKVLQENMRGLLRGSNQSWLKKFQQELVKVCEKLKEENKGSFSNDELFVLCVQPA
KAARLLTHDLRMTTIFLRQQLDQKRDFTVKTAKGIREKCDLGKGFQKEITWPCHTL
EQQCNRGLGTTEILKQVLLNEHKDTLKTHENCVTYLKEKCNKWSRRGDRFSFVCVFQN
ATCKLMVKDVQDRCKIFKENIKVSEIVDFLKNNTNNITTILERNCPSWHTYCNRFSSNC
PDFSKKNPCTKIKNNCKPFYERKALEDALKVELRGKLSDENKCTAALKGYCTLAGNVN
NASVRSLCKDNTQGSNKKTDEKVVEELCKKLMEEVKEQCETLPAELKQPADDLEKDVK
TYEELKEEAKKAMNKSSLVLSFVKKDGNTPKNNSKSEDKNVVSNEKDTIKHVKILRR
GVKDVLVTELEAKAFDLAAEVFGRYVDLKERCEKLTLDGIGKDDCDGLKGVCVGKIKKK
CRDLKPLEVKSHIEIVTESTTTTTTVDPKATECKSLQTTDTWVTQTSTHTSTST
ITSTITSKITLTSTRRKPTKCTTGDDAEDVKPSEGDRVSGWNVMRGAIVAMVISFMI
"

BASE COUNT 1312 a 505 c 753 g 793 t

ORIGIN

Query Match 57.3%; Score 85.4; DB 8; Length 3363;
Best Local Similarity 78.5%; Pred. No. 4.5e-14;
Matches 117; Conservative 0; Mismatches 26; Indels 6; Gaps 1;

Qy 1 tgacctcgacgaggcggtgtaagcctacaaggatgtacgacagagaggaatgaaggcag 60
|||| |||| |||||||| |||| |||| |||||||||||| ||||| ||||
Db 3099 TGACATCAACAAGGCGGTGCAAACCAACCAAGTGTACGACAGG-----GGATGATGCAG 3152

Qy 61 gagaggtgaagccgagtgaggggctgaggatgagtgggtggagtgtgatgagaggggtgt 120
||| ||||||| ||||| ||| ||||| ||||| ||||||| |||||
Db 3153 AAGACGTGAAGCCAAGTGAAGGCTTGAGGGTGAGCGGGTGGAATGTGATGAGGGGGCAA 3212

Qy 121 tattagcaatgatgatttcattcatgatt 149
|| |||||||| | ||||| |||||||||
Db 3213 TAGTAGCAATGGTTATTCGTTCATGATT 3241

SEQ ID NO: 5, 2845-3090 nucleotides
RESULT 1
AF033208
LOCUS AF033208 3089 bp DNA PLN 10-SEP-1998
DEFINITION Pneumocystis carinii f. sp. hominis clone HUMSG11 major surface glycoprotein (MSG) gene, partial cds.
ACCESSION AF033208
VERSION AF033208.1 GI:3560512
KEYWORDS .
SOURCE Pneumocystis carinii f. sp. hominis.
ORGANISM Pneumocystis carinii f. sp. hominis
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; Pneumocystis.
REFERENCE 1 (bases 1 to 3089)
AUTHORS Garbe,T.R. and Stringer,J.R.
TITLE Molecular characterization of clustered variants of genes encoding major surface antigens of human Pneumocystis carinii
JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)
MEDLINE 94314421
REFERENCE 2 (bases 1 to 3089)
AUTHORS Mei,Q., Turner,R.E., Sorial,V., Klivington,D., Angus,C.W. and Kovacs,J.A.
TITLE Characterization of major surface glycoprotein genes of human Pneumocystis carinii and high-level expression of a conserved region
JOURNAL Infect. Immun. 66 (9), 4268-4273 (1998)
MEDLINE 98380374
REFERENCE 3 (bases 1 to 3089)
AUTHORS Mei,Q., Turner,R., Sorial,V., Klivington,D., Angus,C.W. and Kovacs,J.A.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1997) Critical Care Medicine Dept., National Institutes of Health, 10 Center Drive, MSC 1662, Bethesda, MD 20892-1662, USA
FEATURES Location/Qualifiers
source 1. .3089
/organism="Pneumocystis carinii f. sp. hominis"
/sub_species="hominis"
/specific_host="Homo sapiens"
/db_xref="taxon:42068"
/clone="HUMSG11"
/note="derived from HIV-infected human with P. carinii pneumonia"
gene <1. .3089
/gene="MSG"
CDS <1. .3089
/gene="MSG"
/note="surface antigen"
/codon_start=3
/product="major surface glycoprotein"
/protein_id="AAC34971.1"
/db_xref="GI:3560513"
/translation="ARAVKRRAKGAQNSIDEEHVLALILKKNGLEDTKCKTKLEEYCK
TLTNAGLNPEVKHEKLKDPCDNGKRNEKCQDLKVNQKCIKFQGKLQTAAGKKISEL
TDEDCKKNEQQCLFLEGACPTELKDDCNKLRRNNCYQKERNNVAEVLLRALRGDLNET
KTCEKKLKEVCPKLERESDELTELCLYQKTTCVSLVTKGSKCDTLEKEVEEALKNE
LREKCLLLEQCYFHRGNCEGDKSKCNKPNNKDCKEYVPECDELAEKCGKENIVYMHP
GSDFDPTKPEPTLAEDIGLEELYKRAEEDGI FVGRQHVRDATAALLALLKTLKKEEC
IKALKKNCENPHEHEALENLCKENKPSSDGTKCDELEKDVNKTCTSLTSTILKNRLY
ISP D GIAEWGKLPTFLSDEDCAKLESYCFYYKETCPDVKEACMNVR AACYKRG DARA
NSVLQKNMRG L LHGSNKDWLKKFQQELAKVCEKLKG NKG FS NDELFVLCI QPAKAAR
LLTHHHQMRVIFLRQQLDQKRDFPTDKDC KELGRK CQDLG KDS KEITWPCH TLEQQCN
RLGITEILKQILLDEHKDTLKSHENCAYKLKRKCHKWSRRGDDRFSFVCVFQNATCEL

MVKDVQDRCKI FEENMQASDINDSLKKNQIKAESAANICPSWHPYCDRFLPNCPDLKK
GKTFCQNLKYCEPFYKRKVLEDALKVELRGNLNSITKCEPALERYCTVLKDVNNA
SSLCKDNTESKTKADNKVRKKLCLKLVEEVEQQCKVLPTELTELEKSLKKDVKTYE
ELKERAKKAMKSSLVLSLVKKNESNTSKNNSKNKNVVSNGLQDTTKYVKILRRGV
KEALVTESEAKAFDLAAEVFGRYVDLKEKCEKLTSDCGIKDDCDGLKEVCVGKIEKTCH
DLKPLEVKHSHEIVTESTTTTTTVDPKATECKSLQTTDTWVTQTSTHTSTTIT
STITSKITLTSTRRKPTKCTTGDEAGDVKPSEGALKMSGWSVMRGVI VAMVISFMI "

BASE COUNT 1254 a 447 c 662 g 726 t
ORIGIN

Query Match 100.0%; Score 246; DB 8; Length 3089;
Best Local Similarity 100.0%; Pred. No. 1.3e-57;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gaatgcaaattccttacagacaacagatacatgggttacacagacatcgacacacacaaggc 60
||| |||||
Db 2844 GAATGCAAATCCTTACAGACAACAGATACATGGTTACACAGACATCGACACACACAAGC 2903

Qy 61 acgtctaccatcacatctaccatcacatcaaaaataacattgacatcaacgaggcgatgc 120
||| |||||
Db 2904 ACGTCTACCATCACATCTACCATCACATCAAAAATAAACATTGACATCAACGAGGGCGATGC 2963

Qy 121 aaacccaaccaagtgtacgacaggggatgaaggcaggagacgtgaaaccgagtggattg 180
||| |||||
Db 2964 AAACCCAACCAAGTGTACGACAGGGGATGAAGCAGGAGACGTGAAACCGAGTGAGGGATTG 3023

Qy 181 aagatgagtgggtggagcgtgatgaggggggtgatagtagcaatggttatTCGTTATG 240
||| |||||
Db 3024 AAGATGAGTGGGTGGAGCGTGATGAGGGGGTGATAGTAGCAATGGTTATTCGTTATG 3083

Qy 241 atttag 246
||| |||
Db 3084 ATTTAG 3089

SEQ ID NO: 5, nucleotides 2845-3090

RESULT 6

PMCMMSGI

LOCUS PMCMMSGI 3363 bp DNA PLN 26-SEP-1994
DEFINITION Pneumocystis carinii B-cell receptor (msgI) gene, 3' end.
ACCESSION L27092
VERSION L27092.1 GI:535706
KEYWORDS B-cell receptor.
SOURCE Pneumocystis carinii DNA.
ORGANISM Pneumocystis carinii
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; Pneumocystis.
REFERENCE 1 (bases 1 to 3363)
AUTHORS Garbe, T.R. and Stringer, J.R.
TITLE Molecular characterization of clustered variants of genes encoding
major surface antigens of human Pneumocystis carinii
JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)
MEDLINE 94314421
FEATURES Location/Qualifiers
source 1. .3363
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"
mat_peptide 152. .3241
/gene="msgI"
/standard_name="major surface antigen"
/evidence=experimental
/product="B-cell receptor"
gene 152. .3244
/gene="msgI"
CDS <152. .3244
/gene="msgI"
/codon_start=1
/product="B-cell receptor"
/protein_id="AAA21645.1"
/db_xref="GI:535707"
/translation="VARAVKRQVAGVKNNAEERLFALIMRADYKDESKCKNKKIKEYC
DGLKNASLTSEEVHKELEKDFCKDGSGKKCEELKKNEAKCNNFKTKLEGVLVKKDASG
LTNDCKENERQCLFLEGACPDIVEDCSKLRLCYQKKREGVAEEVLLRALRGDLGNK
TECEKKIKIDVCPKIGQESDELLCLDQKKCTNLMTARDKKCNTLEEDVKKALEKNKN
NLLGKCLPLLEHATFTEGTAKKASQCTPNKDCEDYLPKCDELAECGKKGIIYIHPGP
DFDPTKPEPTVAEDIGLEELYKKAEDGVHIGKPPVRDATALLLIQNPDPKIQANE
KEKCKKVLENKCKELKKHEVLGDLCNQNAASQSGTKKCEELEKELANSTKILSEKIKN
KHLGSGSETIPWYKLSTFLSDSDCARLESCDFYFAQDKDPLKKECKNVKAACYKRGKD
ARANKVLQENMRGLLRGSNQSWLKKFQQELVKVCEKLKEENKGSFSNDELFLVCVQPA
KAARLLTHDLRMTTIFLRQQLDQKDFPTVKTAKGIREKQCQDLGKGFQKEITWPCHTL
EQOCNRGTTIELKQVLLNEHKDTLKTHENCVTYLKEKCNKWSRRGDRFSFVCVFQN
ATCKLMVKDVQDRCKIFKENIKVSEIVDFLKNNTNNITLERNCPSWHTYCNRFSSNC
PDFSKKNPCTKIKNNCKPFYERKALEDALKVELRGKLSDENKCTAAALKGYCTLGNVN
NASVRSLSCKDNTQGSNKKTDEKVVVEELCKLMEEVKEQCETLPAELKQPADDLEKDV
TYEELKEEAKKAMNKSSLVLSFVKKDGNNTPKNNSKSEDKNVVSNEKDTIKHVKILRR
GVKDVLVTELEAKAFDLAAEVFGRYVDLKERCEKLTLDGCIKDDCDGLKGVCVGKIKKK
CRDLKPLEVKSHIEVTESTTTTTTVTDPKATECKSLQTTDTWVTQTSTHTSTST
ITSTITSKILTSTRRCKPTKCTGDDAEDVPKSEGRLVSGWNVMRGAIVAMVI SFM
"

BASE COUNT 1312 a 505 c 753 g 793 t
ORIGIN

LOCUS PCU39660 629 bp mRNA PL 29-NOV-1995
 DEFINITION **Pneumocystis carinii major surface glycoprotein (msg) mRNA, partial cds.**
 ACCESSION U39660
 NID g1079706
 KEYWORDS .
 SOURCE **Pneumocystis carinii.**
 ORGANISM **Pneumocystis carinii**
 Eukaryotae; mitochondrial eukaryotes; Fungi; Ascomycota;
 Archaeascomycetes; Pneumocystidaceae; Pneumocystis.
 REFERENCE 1 (bases 1 to 629)
 AUTHORS Kovacs, J.A., Edman, J.C. and Angus, C.W.
 TITLE Direct Submission
 JOURNAL Submitted (30-OCT-1995) J.A. Kovacs, CCMD, NIH, Building 10, Rm
 7D43, MSC 1662, Bethesda, MD 20892-1662, USA
 FEATURES Location/Qualifiers
 source 1..629
 /organism="Pneumocystis carinii"
 /db_xref="taxon:4754"
 gene 55..629
 /gene="msg"
 CDS 55..>629
 /gene="msg"
 /codon_start=1
 /product="major surface glycoprotein"
 /db_xref="PID:g1079707"
 /translation="MRIAFFALFAQLSCILVYSIAERDFMSLDEIYGGDISFDHEKLE
 FNEYNQVLQMILEKAKKLGTGFVRTKDFSNRRYEGRIELNHLGRRPGVDYFRKGKDVF
 TDGYPRGGHLIEDELSEEAMARPVKRQAVQGAQDEIDEKHLLAFIVKDKYKEEQKCK
 EELEKYCKELKEADKNLENVDDKVKGLCDDK"
 misc_difference 228..234
 /gene="msg"
 /note="in some cDNA clones, there are 6 A's at this
 position, instead of 7. This appears to be due to an error
 in reverse transcription"
 /replace="aaaaaa"
 BASE COUNT 218 a 73 c 154 g 184 t
 ORIGIN
 1 ggaaaatata tttttcttg atatccgtct cgtcttcagt ttgttgtgc aataatgagg
 61 attgcatttt ttgcgccttt tgcgcaactt agttgtattt tagtttattc aatagcagaa
 121 agggatttca tgtcattaga taaaatataat ggaggcgata taagtttga tcataaaaaaa
 181 ctcgaattta acgaatataaa tcaagttta caaatgcttg aaaaggcaaa aaaattggga
 241 accggctttg ttgatagaac caaagatttt tctaataagac gatataagg gagaattggag
 301 ttaaatcatt tggggagacg cccaggagtc gactattttt ggaaagggtgg ggatgttttt
 361 actgtatggtt atcctcgtgg aggtcatttg atcgaggatg agttgtccga agaggcggca
 421 atggcacggc cggtaagag gcaagcgtt caaggagcac aagatgagat tgatgagaaa
 481 cacctttgg ctttcattgt gaaggacaaa tataaagaag aacaaaaatg caaagaagaaa
 541 ctcgagaaat attgtaaaga gttgaaggaa gcagataaaa atctagagaa tgtggatgt
 601 aaagttaaag gactttgtga tgataaaaa
 //

LOCUS AF043102 4002 bp mRNA PLN 06-JUN-1998
 DEFINITION **Pneumocystis carinii surface glycoprotein A (gpA) mRNA, complete cds.**
 ACCESSION AF043102
 NID g3184385
 KEYWORDS
 SOURCE **Pneumocystis carinii.**
 ORGANISM **Pneumocystis carinii**
Eukaryota; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae;
Pneumocystis.
 REFERENCE 1 (bases 1 to 4002)
 AUTHORS Haidaris, C.G., Medzihradsky, O.F., Gigliotti, F. and Simpson-Haidaris, P.J.
 TITLE Molecular characterization of mouse **Pneumocystis carinii** surface glycoprotein A
 JOURNAL DNA Res. 5, 77-85 (1998)
 REFERENCE 2 (bases 1 to 4002)
 AUTHORS Haidaris, C.G., Medzihradsky, O.F., Gigliotti, F. and Simpson-Haidaris, P.J.
 TITLE Direct Submission
 JOURNAL Submitted (14-JAN-1998) Microbiol/Immunol, U. Rochester Med. Ctr., Box 672, 601 Elmwood Ave., Rochester, NY 14642, USA
 FEATURES Location/Qualifiers
 source 1..4002
 /organism="Pneumocystis carinii"
 /db_xref="taxon:4754"
 /lab_host="C.B-17 SCID mouse"
 gene 1..4002
 /gene="gpA"
 CDS 52..3849
 /gene="gpA"
 /codon_start=1
 /product="surface glycoprotein A"
 /db_xref="PID:g3184386"
 /translation="MRIAFFALFAQLSYVLGSFLRDRDLPNEEDVYGYENFGLDPNSPEPSEFLNIEMLRNAKRLGSGKINQGQFLSNRKARRDFDLCRSCNRPGVDYFRKSDYDGGFSEDFSSEDYSQDKRWVEEVAQEAAAMAQPVKRQAAGQAAGNDEIKEEQVQLGLIVKSGYNNDNKCKANLKYCEELKKIDGKLESVDVKVKGLCENGKEGEKCKELKKLETELGAFKTEVENALNNLTDEKCRKYEEKCLLLEGDPNNLEEKCVKLDRDCYRQRQGVAKEILLRALEGKVNNKDECKKRMKEICQGLSEYSDELVFSCFNSDKTCEYLQKNHGDSCCKPLEKELEDKELVKEKCQEYLEKCYFYGSSCKDTKCDKVNPKCKGKIEYEGLDFSPVREKPRFPEKIEVENLYKKEAKGIIVGKPKYKTLRDLALLLIKERNKGDEGEKCKKALEDCESFKHLDYGLEELCGDKDKEDRCKELVEVEDRCTMFNLKELYLKGLSTEFKDDES DYFSWGQVSKLVSMEDCIKFESCFHLERCTNKIGKACENVRVACYKKQDRVLNRYFQEGLKGLIGDLELVTNLEKCKQSVVGNYTKLKEDRRTFTKCHLPTKLCYELLDDVILQSEELEVVLNLRDFPRKEDCVELKKKCKDLESDSYLNHEKCDTLNRRCEYLKVTEELRKRLLKRGDDALRTQGNCTAVLKECEELSFRKGDEFSVSCALREETCSFMVEQTENECLFLKNNMENGKIINKIEKGNETLVEELCTLFDPYCHQYIENCPCDRLLKKASNNSNKGVCLELEEKCKPFFEKLKLENELTHELKGSLDKDECKKALRKHCSEQKNSANQKFNSLCNTDKDKDVEEKVCKKLVEKVKRKCPITLENLNEEKNELKKKDEYEAKQSEKFTKEAKLVLRSRPEQDGQGGGSKAQDGSPVPKPVGPPVQPPAPAQPTPGGVAPTLAPPAPQPTSGGAPLPVPPAAPAAPGAPSTPGTPAAPAGPAAPGTPSTPSTPPAGPAGPSGGTPGAPAGPPAPGGSTPSGTTNTSNTVILVRRFTVSGEVSEPEKKAFVATARALELYLELKEKCKGLKGDCERFKDCPKCETVCKEIDELCEGIEGLKVTPHHTVTSTATQTTTTATTTTTTTTTTTTTTTATTESVDGGKVTEECTLVOTTDTWVTSTSHTSTLTSTVTSTVTLTSMRKCKPTRCTSDDSSKETETQKEEEKEEVKPNEGMKIRVPEMIKIMLLGVIVM
 GML"

BASE COUNT 1616 a 487 c 940 g 959 t
 ORIGIN

```

1 caaaaatata ttttcttga tatccgtctc ttttctcaaa tttgtgcaat tatgaggatt
61 gcatttttg cgcttttg cacaacttagt tatgtcttgg ggagttttt aagagataga
121 gatttgccaa atgaggaaga tgtttatggt tatgaaaatt ttgggttggta tcccaattca
181 ccggaaccga gtgaatttt aaatattata gaaatgctc gaaatgc当地 gcgatttagga
241 agtggaaaga ttaaccaggc tcaatttctc tcgaatcgta aggcttaggag agattttgat
  
```

LOCUS D31920 81 bp mRNA PL 22-SEP-1997
DEFINITION **Pneumocystis carinii** MSG mRNA for major surface glycoprotein,
5'UTR.
ACCESSION D31920
NID g2443384
KEYWORDS MSG.
SOURCE **Pneumocystis carinii** cDNA to mRNA.
ORGANISM Pneumocystis carinii
Eukaryotae; Fungi; Ascomycota; Archaeascomycetes;
Pneumocystidaceae; Pneumocystis.
REFERENCE 1 (sites)
AUTHORS Wada,M., Sunkin,S.M., Stringer,J.R. and Nakamura,Y.
TITLE Antigenic variation by positional control of major surface
glycoprotein gene expression in *Pneumocystis carinii*
JOURNAL J. Infect. Dis. 171 (6), 1563-1568 (1995)
MEDLINE 95287050
REFERENCE 2 (bases 1 to 81)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1994) to the DDBJ/EMBL/GenBank databases.
Yoshikazu Nakamura, Institute of Medical Science, University of
Tokyo, Tumor Biology; Shirokanedai 4-6-1, Minato-ku, Tokyo 108,
Japan (E-mail:nak@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5307,
Fax:03-5449-5415)
FEATURES Location/Qualifiers
source 1..81
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"
gene 1..81
/gene="MSG"
5'UTR 1..81
/gene="MSG"
/note="major surface glycoprotein"
BASE COUNT 30 a 10 c 13 g 28 t
ORIGIN
1 attcatgaag taactatcaa gcttaataaa ggacccaata taggagtata ttgttaatat
61 ttatgtata tattggatc c
//

LOCUS D31914 496 bp mRNA PL 22-SEP-1997
DEFINITION *Pneumocystis carinii* MSG mRNA for major surface glycoprotein,
5'UTR.
ACCESSION D31914
NID g2443378
KEYWORDS MSG.
SOURCE *Pneumocystis carinii* cDNA to mRNA.
ORGANISM *Pneumocystis carinii*
Eukaryotae; Fungi; Ascomycota; Archaeascomycetes;
Pneumocystidaceae; *Pneumocystis*.
REFERENCE 1 (sites)
AUTHORS Wada,M., Sunkin,S.M., Stringer,J.R. and Nakamura,Y.
TITLE Antigenic variation by positional control of major surface
glycoprotein gene expression in *Pneumocystis carinii*
JOURNAL J. Infect. Dis. 171 (6), 1563-1568 (1995)
MEDLINE 95287050
REFERENCE 2 (bases 1 to 496)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1994) to the DDBJ/EMBL/GenBank databases.
Yoshikazu Nakamura, Institute of Medical Science, University of
Tokyo, Tumor Biology; Shirokanedai 4-6-1, Minato-ku, Tokyo 108,
Japan (E-mail:nak@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5307,
Fax:03-5449-5415)
FEATURES Location/Qualifiers
source 1..496
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"
gene 1..496
/gene="MSG"
5'UTR 1..496
/gene="MSG"
/note="major surface glycoprotein"
BASE COUNT 214 a 50 c 89 g 143 t
ORIGIN
1 caactatcaa gcttaataaa ggacccaata taggagtata ttgttaatat ttatgtata
61 tattggatc cagaataat atgaaatgct ttttccaaa attaataaaa ttatagttct
121 aagaatatta attttaattt ttcctactta tggcaatcaa agcacacatt tgaaagccag
181 agattctcaa gtaagcgcag caaatactcc gtttaaggat attgttgttta aggtatgaaga
241 aatccttgct tatattttga gagaagatta taaagataaa tgcgaactaa agcttgagga
301 atattgtaaa gaatttgaggg atatacatca aggattaaat aaagttcata ctatagttaa
361 agaaatttgt gatgtatgaaa aacgagacgg aaaatgcaaa gaactgaaag acaaagttaa
421 aaaggaattt gaaactttt aagaggaact taaaaagaca ttgaaagaca taaaagatga
481 aaatttgtgaa aaatat
//

LOCUS AF035221 1945 bp mRNA PL.. 02-JAN-1998
 DEFINITION **Pneumocystis carinii** glycoprotein A (gpA) mRNA, partial cds.
 ACCESSION AF035226
 NID' g2739185
 KEYWORDS .
 SOURCE *Pneumocystis carinii*.
 ORGANISM *Pneumocystis carinii*
 Eukaryotae; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae;
 Pneumocystis.
 REFERENCE 1 (bases 1 to 1945)
 AUTHORS Guadiz,G., Haidaris,C.G., Maine,G.N. and Simpson-Haidaris,P.J.
 TITLE Direct Submission
 JOURNAL Submitted (19-NOV-1997) Medicine-Vascular Medicine, University of
 Rochester, 601 Elmwood Avenue, Box 610, Rochester, NY 14642, USA
 FEATURES Location/Qualifiers
 source 1..1945
 /organism="Pneumocystis carinii"
 /db_xref="taxon:4754"
 /tissue_type="infected lung"
 /clone="2-1A"
 gene <1..1945
 /gene="gpA"
 CDS <1..1914
 /gene="gpA"
 /codon_start=1
 /product="glycoprotein A"
 /db_xref="PID:g2739186"
 /translation="KKELRKKNFCNDSCSNEKDNEKLVKLRDEDCAKLQSKCFYLE
 KHCSANLKKPCNNVRVTCYKRGKAAAMTLLSVLKKGKLKPGPNNDYKDCQKALLEKC
 KEVRNISAEVFEMCLYPENTCQNLSIDVQQKANYLSLFSLQNRDHPSQEDCVELQEKC
 EALEADADWLRPPCETLRTHCNFLYLSEKLIKHHLLSEGKGKLSSNEICNKELECHS
 WHKKKNETYAFPCALRNESCELMWWRVGKHCDEFKENLKNYNTTLIKNPSEETCLLWA
 PHCENLTPNCEEKVAKDINETCRKLQEKCSPVLEKRNLNKLNRELKGKLTGKTEC
 IDYFKSLCTQADHPNKTALDILCKENGQNIDDDAKEKKCQPLIDEVQWECPLLKTKL
 EKAKDEVAKKADEYKKLKEEAEKAITETKLNVTELPEAGKATLIRRDPKGSSPAKRS
 VKVEVVSVPITKEHAEAIIDLLSQALDLYFELREECDALILDCAFEDCPQCKDSCNSI
 LATCKGLQPLTIAEPRIIEDQKPEAGKEDTPEGGQAGGQAAGKECTSITTSDTWVTST
 STHTTTTTSTKTTVKVTLTSTRCKPTKCTTGKGEPTSRGDEEEVVPSGGRKWLPS
 LSIVIVGVVVVALV"

BASE COUNT 744 a 279 c 451 g 471 t

ORIGIN

```

1 aaaaaaagaac ttctgtaaaaaa gaatttctgc aatgacgata gctgcagtaa tgaaaaagat
61 aattgggaga aatttagtaaa gttaagagat gaagatttg caaaattaca atctaagtgt
121 ttttatatttag aaaaacatgg ttcagctaat cttaaaaaac catgtacaaa tgtagatgt
181 acctgttaca agagaggatt aaaggcagcg gctatgacct tgtagagag tgtagtggaaa
241 gggaaactta aaccagggtcc aaataatgtat tataaagatt gccaaggc attattagaa
301 aagtgtaaag aagtggaaaa tatcagtgc aagtggtt agatgtgtt atatccagaa
361 aatacgtgtc aaaatcttc tattgtatgtt caacaaaaag caaattatgg aagcctttt
421 tcacttcaaa acaggggacca tccaaatgtc aaagactgtg ttgaactaca agagaaatgt
481 gaagcttttag aaggcatgtc agactggctt cgacccctt gtgaaacatt gagaacgcatt
541 tgcaactttc ttacccatgg agaaaaatgg aacatcatt tgtagatgtg agggaaaggt
601 aaattaaatgt gcaatggat ttgcacaaa gagttgtat agagggtgtca ctcatggcat
661 aaaaaaaaaga acgaaacgtt tgctttccg tggctttgc gcaatgtatc ttgtggatgt
721 atgggttggc gtgttagaaaa acatttgtat gaatttaaaag aaaactgtaa aaattataat
781 acaacattaa taaaaatcc atcagaagaa acatgtctt tggccacc acactgtgaa
841 aatctgtacgc caaaactgtgt aaaaaaggtt gctaaagata tcaatggatc ttggccgtaaa
901 ctccaggaga aatgttccacc agttcttgag aagagggatt tgaaaaataa attgaaacgt
961 gagttggaaag gggaaatgttacca ggttggaaag aagacagaat gtatcgatta ttggatgtt
1021 ctttgtacac aggcatgtc tccaaataaa acggcactt acatattatg taaagaaaaat
1081 ggacaaaata ttgtatgtt aaaaatgtt aaaaatgtt gccaaccact cattgtatgtt
1141 gtacaatggg agtgtccact attaaaaatgtt aagttggaaaa aagcaaagga tggatgtggca
1201 aagaaggcag atgaaatataa aaaactaaaaa gaagaggcag aaaaggctat tacggaaaca
1261 aaactcaatgt ttacagaact accaaaggaa gcaggaaagg caaccctgtat tagaagagat
1321 caaaaaggca gcaatgttccacc caagatgttca gttaaaggatgg aggttggatc ttggccgtaaa

```

1381 acgaaaagaac at ggcc aattgattta ttatctcaag ca tct ttatttcgag
1441 ttgagagagg agt, -yatgc gttaatatta gattgtgact tta, ,agga ttgtcctcaa
1501 tgtaaagatt catgtaatag tatacttgct acctgtaaag ggttacaacc tcttacaatc
1561 gcggAACCCAC gtattattga agatcagaaa ccagaAGCTG gaaAGGAAGA tactcctgaa
1621 ggaggacaag ctggaggaca agctcggg aaggagtgc cctctatcac aacaagtgac
1681 acatgggtca cgagtacgtc tacgcatacg acgacgacca cctcgacaaa gacgacgaca
1741 gttaaagtga cgtaaacgtc aacgaggaag tgtaagccga ccaagtgtac cactggaaaa
1801 ggagagccga caagtagagg tgatgaagag gaggagggtt tgccgagtgg aggaaggaag
1861 tggttgccaa gtctaagtat agttattgtt ggtgtgggtt tggcactcgt gtagatcatg
1921 atgaataaga cgttccatgt gaaat

//

LOCUS PCU5392: 2470 bp DNA PL. 24-DEC-1996
 DEFINITION *Pneumocystis carinii* major surface glycoprotein (MSG) gene, partial
 cds.
 ACCESSION U53921
 NID g1750217
 KEYWORDS .
 SOURCE *Pneumocystis carinii*.
 ORGANISM *Pneumocystis carinii*
 Eukaryotae; mitochondrial eukaryotes; Fungi; Ascomycota;
 Archaeascomycetes; Pneumocystidaceae; Pneumocystis.
 REFERENCE 1 (bases 1 to 2470)
 AUTHORS Edman, J.C., Hatton, T.W., Nam, M., Turner, R., Mei, Q., Angus, C.W. and
 Kovacs, J.A.
 TITLE A single expression site with a conserved leader sequence regulates
 variation of expression of the *Pneumocystis carinii* family of major
 surface glycoprotein genes
 JOURNAL DNA Cell Biol. 15 (11), 989-999 (1996)
 MEDLINE 97101113
 REFERENCE 2 (bases 1 to 2470)
 AUTHORS Kovacs, J.A.
 TITLE Direct Submission
 JOURNAL Submitted (09-APR-1996) Joseph A. Kovacs, CCMD/CC, NIH, Bldg. 10,
 RM. 7D43, MSC1662, Bethesda, MD 20892-1662, USA
 FEATURES Location/Qualifiers
 source 1..2470
 /organism="Pneumocystis carinii"
 /db_xref="taxon:4754"
 /lab_host="rat"
 gene 1940..2470
 /gene="MSG"
 CDS join(<1940..2036,2187..>2470)
 /gene="MSG"
 /note="gpA; expression site of MSG; encodes leader
 sequence of MSG; surface protein"
 /codon_start=1
 /product="major surface glycoprotein"
 /db_xref="PID:g1750218"
 /translation="MRIAFFALFAQLSCILVYSIAERDFMSLDEIYEGGDISFDHEKL
 EFNEYNQVLQMLEKAKLGTGFVDRTKDFSNRRYEGRIELNHLGRRPGVDYFRKGDV
 FTDGYPRGGHLLIEDELSEEAMARP"

BASE COUNT 707 a 335 c 411 g 1017 t
 ORIGIN

```

1 gaattccata ataattgtat tgtaaattat gaatataatt ttgatttattt cttaaaaata
61 tagtattttt ttataataat atttctatt atcaacttga tgttaaaaat atcgtaattt
121 atcttgatgt ttagataaa tgtagtattt tcacagtctt attaggtatc cgttggaaatt
181 atttttaaat tgaaaaggta taaagtatga ataataatata atatggggaa ttttattttg
241 tagtatataa atcatagcaa catgatttt tttgtattt aaatcagatt ttataaaaatg
301 tttttatata attataaaat cttgttataag tatatattt caagtgtttt caatgggtt
361 agttttaaag agagtgttaa ttaaggtaga aatgtttat taatgggtat atatatataat
421 atattgtttt ttgtctaata aatgaatat tttctaaaat aaagattatt tttttaaaaa
481 attgattaaa atgtagaatg tatgattatt tcaagtgtt gtatagttact agtcccttgt
541 cgtccctcgtc gcctctgtcg tcctctgtcg ggtgtgtat agagttttat tgaatcatgt
601 attttaaaaa aaataaaagat tatttatata tatataatatt tgaaaaaaatt gaaaggagaa
661 ttatttagta tgatttttaa attttacaag agttgaaatg tttgttgtat ttttatttca
721 ttatataagg ttatgtggaa cgaagttggaa agtttaatatt atgttgaatg agaatatata
781 gggtaggtc gaacgttttag gttttaagaag aatataatagg gttacgtcga aggttttagt
841 tgaagtagaa tatataagggt taggtctaattt gtttaggtta ggtcaaagtg cagggttagg
901 tctaacatcc ggggttaggt caaagtatag tttaggtcga cgtgttagtt aggtctgagt
961 tatgtatagg ttatgtggta gtttcaactt tttttaatg aaaaaacttt ttttaataaa
1021 aaaagtaata aactcaaattt ctatagatta acatggatcc gtttttattttaaaaatttca
1081 tgtttatgt ataaaactttt cttttttta ttttacttg gtccctttct ttccctcatc
1141 atactggaa ttttctgata tattacaata gatgggttcat tcctctactc acttttaaaa
1201 tcttatttaa ctgaccatc atgttatttt tccccttattt tttataaaaaa gactcgtcgc
1261 ttatataatag aatggcgtat acgtaaaaat agaaacgtgt atgcgtgaaa tagaatggtg

```

1321 cttttaaaatt taa .ttt attcaaatta ggaaatctt ttt .ttt tagtttgga
1381 atccattatt ctt .aca aagtatttct tttttccta aca stat ttataacttta
1441 tatttggaaa cctatgcata tactcgatat cactctgact ttgcataattc tactccctcc
1501 ttccctttca cccccctctgg gacactcccc tgtccggaca cacccggac agctcttggc
1561 tgccctccc tcgacccttc tataccatc tatccatcta tccatctata tatacacccgt
1621 ttttactcat ttctatccta taaaaattga cttgtaattt tggatatatt ccttctataat
1681 gcatttttag gactggctaa tattttagg tctagctctt ttatttccg gctagtcattc
1741 cccccatccg gaaattggct ttccctttt cagcggtttc ttgcattcat ccgtaaattt
1801 ccgtcaaatat ttaagggtct ttgcattatta agggactgt ttctgactg tatttttacc
1861 atacccttg ctccctttag cgcttggaaa atatatttt tcttgatatc cgtctcgct
1921 tcagttgtt tggcaataa tgaggattgc atttttgcg cttttgcgc aacttagtt
1981 tatttttagtt tattcaatag cagaaaggga ttcatgtca ttagatgaaa tatatggta
2041 gtttttttc ctaatttcc ttgttatca ttttttacac gtctttttt actcgcttt
2101 tactcatctt ttactcate ttcttattct ctccctttct tttggaaatc ttttttccct
2161 tttatggttt tactcatgtt ttctagaagg aggcgatata agttttgatc ataaaaaact
2221 cgaatttaac gaatataatc aagttttaca aatgcttcaa aaggcaaaaa aattgggaac
2281 cggcttggat gatagaacca aagattttc taatagacga tatgaaggga gaatttgagt
2341 aaatcatttg gggagacgcc caggagtgcg ctattttagg aaagggtgggg atgtttttac
2401 tgatggttat cctcggtggag gtcatttcat cgaggatgag ttgtccgaag aggccggcaat
2461 ggcacggccg

//

LOCUS AF03522. 1945 bp mRNA **PR.** 02-JAN-1998
DEFINITION *Pneumocystis carinii* glycoprotein A (gpA) mRNA, partial cds.
ACCESSION AF035226
NID g2739185
KEYWORDS
SOURCE *Pneumocystis carinii*.
ORGANISM *Pneumocystis carinii*
Eukaryotae; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae;
Pneumocystis.
REFERENCE 1 (bases 1 to 1945)
AUTHORS Guadiz, G., Haidaris, C.G., Maine, G.N. and Simpson-Haidaris, P.J.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1997) Medicine-Vascular Medicine, University of Rochester, 601 Elmwood Avenue, Box 610, Rochester, NY 14642, USA
FEATURES Location/Qualifiers
source 1..1945
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"
/tissue_type="infected lung"
/clone="2-1A"
gene <1..1945
/gene="gpA"
CDS <1..1914
/gene="gpA"
/codon_start=1
/product="glycoprotein A"
/db_xref="PID:g2739186"
/translation="KKELRKKNCNDSCSNEKDWEKLVKLRDECAKLOSKCFYLE
KHCSANLKKPCNNVRVTCYKRGLKAAAMTLLESVLKGKLPGPNNDYKDCQKALLEKC
KEVRNISAEVFEMCLYPENTCQNLSIDVQQKANYLSLFSLQNRDHPSQEDCVELQEKC
EALEADADWLRPPCETLRTHCNFLYSEKLKHILLSEGKGKLISSNEICNKELNERCHS
WHKKKNETYAFFCALRNESCELMWWRVGKHCDFENLKNYNTTLIKNPSEETCLLWA
PHCENLTPNCEEKVAKDINETCRKLQEKCPVLEKRNLLKNLKRELKGKLTGKKT
IDYFKSLCTQADHPNKTALDILCKENGQNIDDDKAEEKKCQPLIDEVQWECPLLKTKL
EKAKDEVAKKADEYKKLKEAEKAITETKLNVTLPKEAGKATLIRRDPKGSSPAKRS
VKVEVVSVVPITKEHAEADLLSQALDLYFELRECDALILDCDFKEDCPQCKDSCNSI
LATCKGLQPLTIAEPRIIEDQKPEAGKEDTPEGGQAGGQAAGKECTSITTSDTWVTST
STHTTTTSTKTTVKTILTSTRCKPTKCTTGKGEPTSRGDEEEEVVPSGGRKWLPS
LSIVIVGVVVALV"
BASE COUNT 744 a 279 c 451 g 471 t
ORIGIN

```

1 aaaaaagaac ttctgtaaaaaa gaatttctgc aatgacgata gctgcagtaa tgaaaaagat
61 aattgggaga aatttagtaaa gtttagagat gaagatttg caaaattaca atctaagtgt
121 ttttatttag aaaaacattt gttcgtcaat cttaaaaaac catgtaccaa tgttagagta
181 acctgttaca agagaggatt aaaggcagcg gctatgactt tgtagagag tggttgaaa
241 gggaaactta aaccaggccc aaataatgat tataaagatt gccaaaaagc attattagaa
301 aagtgtaaag aagtggagaa tatcagtgcga gaagtgttt agatgtgtt atatccagaa
361 aatacgtgtc aaaatcttc tattgtatgtt caacaaaaag caaattat t aagccctttt
421 tcacttcaaa acaggccca tccaaagtccaa gaagactgtg ttgaactaca agagaaatgt
481 gaagcttttag aagcagatgc agactggctt cgacccctt gtgaaacatt gagaacgcatt
541 tgcaactttc ttacccccc agaaaaggttt aaacatcatt tgtaggtga aggaaaggt
601 aaattaagta gcaatgagat ttgcaacaaa gagttgaatg agaggtgtca ctcatggcat
661 aaaaaaaaaga acgaaacgtt tgctttccg tggctttgc gcaatgaaatc ttgtgagttt
721 atggtttgc gtgttagaaaa acattgttat gatattaaag aaaacttggaa aaattataat
781 acaacattaa ttaaaaatcc atcagaagaa acatgtctt tggggcacc acactgtgaa
841 aatctgacgc caaactgtga agaaaagggtt gctaaagata tcaatgagac ttggccgtaaa
901 ctccaggaga aatgttccacc agttcttgag aagaggaaat tgaaaaataa attggaaacgt
961 gagttgaaag gggaaatataac ggatggaaag aagacagaat gtatcgat tttttaaaagtt
1021 ctttgtacac aggccatca tccaaataaa acggcactt acatattatg taaaagaaaaat
1081 ggacaaaata ttgtatgtat taaagctaaa gaaaagaaaaat gccaaccact cattgtatgaa
1141 gtacaatggg agtgtccact attaaaaaca aagttgaaaa aagcaaaagg tggaaatggca
1201 aagaaggcag atgaatataa aaaactaaaaa gaagaggcag aaaaggctat tacggaaaca
1261 aaactcaatg ttacagaact accaaaggaa gcaggaaagg caaccctgtat tagaagagat
1321 caaaaaggca gcagtccagc caagagatca gttaaaggtt aggttggtc tggccatata

```

1381 acgaaagaac atc aggc aattgattta ttatctcaag cat itct ttatttcgag
1441 ttgagagagg agtgytatgc gtaaatatta gatttgact tta agga ttgtcctcaa
1501 tgtaaagatt catgtaatacg tatacttgct acctgtaaag gtttacaacc tcttacaatc
1561 gcggaaaccac gtattattga agatcagaaa ccagaagctg gaaaggaaga tactcctgaa
1621 ggaggacaag ctggaggaca agctgcggg aaggagtgc cctctatcac aacaagtgc
1681 acatgggtca cgagtacgtc tacgcatacg acgacgacca cctcgacaaa gacgacgaca
1741 gttaaagtga cgttaacgtc aacgaggaag tgtaagccga ccaagtgtac cactggaaaa
1801 ggagagccga caagtagagg tgatgaagag gaggaggtgg tgccgagtttggactcgt gtagatcatg
1861 tgttgc当地 gtctaagtat agttattgtt ggtgttgtt ggactcgt gtagatcatg
1921 atgaataaga cgttccatgt gaaat

//

LOCUS PCU53921 **2470 bp** **DNA** **PLN.** **24-DEC-1996**
DEFINITION *Pneumocystis carinii* major surface glycoprotein (MSG) gene, partial
cds.
ACCESSION U53921
NID g1750217
KEYWORDS .
SOURCE *Pneumocystis carinii*.
ORGANISM *Pneumocystis carinii*
Eukaryotae; mitochondrial eukaryotes; Fungi; Ascomycota;
Archaeascomycetes; Pneumocystidaceae; Pneumocystis.
REFERENCE 1 (bases 1 to 2470)
AUTHORS Edman, J.C., Hatton, T.W., Nam, M., Turner, R., Mei, Q., Angus, C.W. and Kovacs, J.A.
TITLE A single expression site with a conserved leader sequence regulates variation of expression of the *Pneumocystis carinii* family of major surface glycoprotein genes
JOURNAL DNA Cell Biol. 15 (11), 989-999 (1996)
MEDLINE 97101113
REFERENCE 2 (bases 1 to 2470)
AUTHORS Kovacs, J.A.
TITLE Direct Submission
JOURNAL Submitted (09-APR-1996) Joseph A. Kovacs, CCMD/CC, NIH, Bldg. 10, RM. 7D43, MSC1662, Bethesda, MD 20892-1662, USA
FEATURES Location/Qualifiers
source 1..2470
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"
/lab_host="rat"
gene 1940..2470
/gene="MSG"
CDS join(<1940..2036,2187..>2470)
/gene="MSG"
/note="gpA; expression site of MSG; encodes leader sequence of MSG; surface protein"
/codon_start=1
/product="major surface glycoprotein"
/db_xref="PID:g1750218"
/translation="MRIAFFALFAQLSCILVYSIAERDFMSLDEIYEGGDISFDHEKL
EFNEYNQVLQMILEKAKKLGTGFVDRTKDFSNRRYEGRIELNLGRRPGVDYFRKGGDV
FTDGYPRGHHLIEDELSEAAAMARP"

BASE COUNT 707 a 335 c 411 g 1017 t
ORIGIN

```

1 gaattccata ataattgtat tgtaaattat gaatataatt ttgatttatt cttaaaaata
61 tagtattttt ttataataat attttctatt atcaacttga tggaaaaat atcgtaattt
121 atcttgatgt ttagagtaaa tggatatt ttcaagtcgt attaggatc cggtggaaatt
181 attttttaat tgaaaaggta taaagtatga ataataatata atatggggaa ttttattttg
241 tagtatataa atcatagcaa catgattttt tttgttattt aaatcagatt ttataaaaatg
301 ttttttatatttataat cttgtttag tatataattt caagtgtttt caatgggttt
361 agttttaaag agagtgttaa ttaaggtaga aatgtttat taatggtgat atatatata
421 atattgtttt ttgtctaata aatgaatattttttaaaat aaagattatt tttttaaaaaa
481 attgattaaa atgtagaatg tatgattatt tcaagtgtt gtatagtaat agtccctcgat
541 cgtcctcgtc gtcctcgctc tcctcgctg ggttgtat agagtttat tgaatcatgt
601 attttaaaaaa aaataaaagat tattttatata tattttatattt tggaaaaattt gaaaggagaa
661 ttatttttagta tgatttttaa attttacaag agttgaaatg tttgttgtat ttttatttca
721 ttattaaggt ttagttggaa cgaagttgaa agtttaatattt atgttgaatg agaatatata
781 gggtagtgc gaacgttttag gtttaagaag aatataatagg gttacgtcga agtttaggt
841 tgaagtagaa tatataagggt taggtctaat gtttaggtt ggtcaaaatg cagggttagg
901 tctaaccatc ggggttaggt caaagtatag gttaggtcga cgtgttagtt aggtctgagt
961 tatgtatagg gtttagggta gttcactct tttttgaatg aaaaaacttt ttttaataaa
1021 aaaagtaata aaactcaaattt ctatagatta acatggatcc gtttttattttaaaaatttca
1081 tggatattgt ataaaaactttt ctttttttaa tttttacttgc tccctttctt tccctcatc
1141 atacttggaa ttttctgata tattacaata gatggttcat tccctactc acttttaaaa
1201 tcttatttaa ctgaccattc atgctattttt tcccttatttttataaaaaa gactcgctcg
1261 ttatataatag aatggcgtat acgtaaaaat agaaacgtgt atgcgtaaaa tagaatgggt

```

1321 cttttaaattt ta~~t~~ tttt attcaaatta ggaaatcttt tttt tagttttgga
1381 atccattatt ct~~t~~ .taca aagtatttct tt~~t~~tttccta aca stat ttatacttta
1441 tatttgaaa cctatgcata tactcgatat cactctgact ttgcataattc tactccctcc
1501 ttcccttca cccccctctgg gacactcccc tgcggaca caccggac agctttggc
1561 tgccctccc tcgacccttc tatacccatc tatccatcta tccatctata tatacacccgt
1621 tttaactcat ttctatccta taaaaattga cttgtatattt tggatatattt ccttctatatt
1681 gcatttttag gactggctaa tatttttaagg tctagctttt ttatattccg gctagtcatc
1741 ccccatccg gaaatggct ttccctttt cagcggtttc ttgcattcat ccgtaaattt
1801 ccgtcaatat ttaagggtct ttgcattatc agggactgt ttctggactg tatttttacc
1861 atacccgtt cttcccttag cgcttggaaa atatattttt tcttgatatc cgtctcgct
1921 tcagtttgtt tgtcaataa tgaggattgc atttttgcg cttttgcgc aacttagttg
1981 tatttttagtt tattcaataag cagaaaggga ttcatgtca tttagatgaaa tatatggtga
2041 gtttttttc ctttatcc ttgtgtatca ttttttacac gtctttttt actcgcttt
2101 tactcatctt ttactcatc ttcttattct ctcctttct ttggaaatc tcttttccct
2161 ttatggttt tactcatgtt ttctagaagg aggcatata agttttgatc ataaaaact
2221 cgaatttaac gaatataatc aagttttaca aatgcttcaa aaggcaaaaa aattgggaac
2281 cggcttggat gatagaacca aagattttc taatagacga tatgaagggaa gaattgagtt
2341 aaatcattt gggagacgcc caggagtgcg ctatattttagg aaagggtgggg attttttac
2401 tgatggttat cctcgtggag gtcatttgat cgaggatgag ttgtccgaag aggccggcaat
2461 ggcacggccg

//

LOCUS PCU3966 629 bp mRNA P 29-NOV-1995
 DEFINITION **Pneumocystis carinii major surface glycoprotein (msg) mRNA, partial cds.**
 ACCESSION U39660
 NID g1079706
 KEYWORDS .
 SOURCE **Pneumocystis carinii.**
 ORGANISM **Pneumocystis carinii**
 Eukaryotae; mitochondrial eukaryotes; Fungi; Ascomycota;
 Archaeascomycetes; Pneumocystidaceae; Pneumocystis.
 REFERENCE 1 (bases 1 to 629)
 AUTHORS Kovacs, J.A., Edman, J.C. and Angus, C.W.
 TITLE Direct Submission
 JOURNAL Submitted (30-OCT-1995) J.A. Kovacs, CCMD, NIH, Building 10, Rm
 7D43, MSC 1662, Bethesda, MD 20892-1662, USA
 FEATURES Location/Qualifiers
 source 1..629
 /organism="Pneumocystis carinii"
 /db_xref="taxon:4754"
 gene 55..629
 /gene="msg"
 CDS 55..>629
 /gene="msg"
 /codon_start=1
 /product="major surface glycoprotein"
 /db_xref="PID:g1079707"
 /translation="MRIAFFALFAQLSCILVYSIAERDFMSLDEIYGGDISFDHEKLE
 FNEYNQVLQMLEKAKKLGTFVDRTKDFSNRRYEGRIELNHLGRRPGVDYFRKGDDVF
 TDGYPRGGHLIEDELSEAAAMARPVKRQAVQGAQDEIDEKHLLAFIVKDKYKEEQCK
 EELEKYCKELKEADKNLENVDDKVKGLCDDK"
 misc_difference 228..234
 /gene="msg"
 /note="in some cDNA clones, there are 6 A's at this
 position, instead of 7. This appears to be due to an error
 in reverse transcription"
 /replace="aaaaaa"
 BASE COUNT 218 a 73 c 154 g 184 t
 ORIGIN
 1 ggaaaatata ttttttcttg atatccgtct cgtcttcagt ttgtttgtgc aataatgagg
 61 attgcatttt ttgcgtttt tgcgcaactt agttgtattt tagtttattc aatagcagaa
 121 agggatttca tgtcatttga tgaaatatat ggaggcgata taagtttga tcattaaaaaa
 181 ctcgaattta acgaatataa tcaagttta caaatgcttg aaaaggcaaa aaaattggga
 241 accggctttg ttgatagaac caaagatttt tctaataagac gatatgaagg gagaattgag
 301 ttaaatcatt tggggagacg cccaggagtc gactattta ggaaagggtgg ggatgtttt
 361 actgatggtt atccctgtgg aggtcatttgc atcgaggatg agttgtccga agaggcggca
 421 atggcacggc cggttaagag gcaagcagta caaggagcac aagatgagat tgatgagaaa
 481 caccctttgg ctttcatttgt gaaggacaaa tataaagaag aacaaaaatg caaagaagaa
 541 ctcgagaaat attgtaaaga gttgaaggaa gcagataaaa atctagagaa tgtggatgat
 601 aaagttaaag gactttgtga tgataaaaa
 //

LOCUS AF043102 4002 bp mRNA PLN 06-JUN-1998
 DEFINITION **Pneumocystis carinii surface glycoprotein A (gpA) mRNA, complete cds.**
 ACCESSION AF043102
 NID g3184385
 KEYWORDS .
 SOURCE Pneumocystis carinii.
 ORGANISM Pneumocystis carinii
 Eukaryota; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae;
 Pneumocystis.
 REFERENCE 1 (bases 1 to 4002)
 AUTHORS Haidaris,C.G., Medzihradsky,O.F., Gigliotti,F. and Simpson-Haidaris,P.J.
 TITLE Molecular characterization of mouse **Pneumocystis carinii surface glycoprotein A**
 JOURNAL DNA Res. 5, 77-85 (1998)
 REFERENCE 2 (bases 1 to 4002)
 AUTHORS Haidaris,C.G., Medzihradsky,O.F., Gigliotti,F. and Simpson-Haidaris,P.J.
 TITLE Direct Submission
 JOURNAL Submitted (14-JAN-1998) Microbiol/Immunol, U. Rochester Med. Ctr., Box 672, 601 Elmwood Ave., Rochester, NY 14642, USA
 FEATURES Location/Qualifiers
 source 1..4002
 /organism="Pneumocystis carinii"
 /db_xref="taxon:4754"
 /lab_host="C.B-17 SCID mouse"
 gene 1..4002
 /gene="gpA"
 CDS 52..3849
 /gene="gpA"
 /codon_start=1
 /product="surface glycoprotein A"
 /db_xref="PID:g3184386"
 /translation="MRIAFFALFAQLSYVLGSFLRDRDLPNEEDVYGYENFGLDPNSPEPSEFLNIIEMLRNAKRLGSGKINQGQFLSNRKARRDFDLCRSCNPGVDYFRKSDYDGGFSEDFSSEDYSQDKRWVEEVAKQEAAMAQPVKRQAAGQAAGNDEIKEEQVLGLIVKSGYNNDNKCKANLKYCEELKKIDKLESVDVKVKGLCENGKEGEKCKELKKLETELGAFKTEVENALNNLTDEKCRKYEEKCLLLEEGDPNNLEEKCVKLDRDCYRQRQGVAKEILLRALEGKVNNKDECKKRMEICQGLSEYSDELVFSCFNSDKTCEYLQKNHGDSCCKPLEKELEDKELVKEKCQEYLEKCYFYGSSCKDTKCDKVNNKCKGKGIEYECPKLDfspvREKPRFPEKIEVENLYKKEAKGIIVGKPKYKTLRDLALLLIKERNKDEGEKCKKALEDCESFKHLDYGLEELCGDKDKEDRCKELVVEDRCTNFKELYLKGLSTEFEKDKESDYFSWGQVSKLVSMEDCIKFESECFHLERVCTNKIGKACENVRVACYKKGQDRVLNRFQEGLKGLIGDLELVTNLEKCKQSVGNYTKLKEDRRYFTKCHLPTKLCYELLDDVIHQSEELEVVLNLRRDFPRKEDCVELKKKCKDLESDSLHEKCDTLNRRCEYLKVTIELRKRLKRGDDALRTQGNCTAVLKKECEELSKRGKDEFVSICALREETCSFMVEQTENECLFLKNNMENGKIINKIEKGNETLVEELCTLFDPYCHQYIENCPCDRLKKASNSNKNGVCLEEEKCKPFFEKLKLENELTHELGKSLDKDECKKALRKHCSEQKNSANQKFNLCNTDKDKDVEEKVCKKLVEKVKRKCPITLENKLNEEKNELKKKKDEYEKAKQESEKFTKEAKLVLRSRPEQDGQGGGSKAQDGSPVPKPVGPPVQPPAPAQPPTPGGVPAPLITAPAQPTSGGAPLPVPPAAPAAGAPSTPGTPAAPAGPAAPGTPSTPSTPPAGPAGPSGGTPGAPAGPPAPGGSTPSGTTNTSVILVRRTFVSGEVSEPEKKAFVATARALELYLELKEKCKGLKGDCEFRKDCPKCETVCKEIDELCEGIEGLKVTPHHTVTSTATQTTTTATTTTTTTTTTTTTTTTATTTESVDGGKVTEECTLVQTTDTWVTSTSHTSTLTSTSTVTSVTSTVTLTSMRKCKPTRCTSDDSSKETETQKEEEKEEVKPNEMKIRVPEMIKIMLLGVIM
 GML"

BASE COUNT 1616 a 487 c 940 g 959 t
 ORIGIN

```

    1 caaaaatata ttttttttgc tatccgtctc ttttctcaaa tttgtgcaat tatgaggatt
    61 gcatttttttgc cgctttttgc gcaacttagt tatgtcttgg ggagttttt aagagataga
   121 gatggccaa atgaggaaga tgtttatggt tatggaaatt ttgggttggta tcccaattca
   181 ccggaaaccga gtgaattttt aaatattata gaaatgctc gaaatgcaa gcgatttagga
   241 agtgggaaga ttaaccagggt tcaatttctc tcaaatcgta aggcttaggag agattttgat
  
```

LOCUS D31920 81 bp mRNA PL 22-SEP-1997
DEFINITION **Pneumocystis carinii** MSG mRNA for major surface glycoprotein,
5'UTR.
ACCESSION D31920
NID g2443384
KEYWORDS MSG.
SOURCE **Pneumocystis carinii** cDNA to mRNA.
ORGANISM **Pneumocystis carinii**
Eukaryotae; Fungi; Ascomycota; Archaeascomycetes;
Pneumocystidaceae; Pneumocystis.
REFERENCE 1 (sites)
AUTHORS Wada,M., Sunkin,S.M., Stringer,J.R. and Nakamura,Y.
TITLE Antigenic variation by positional control of major surface
glycoprotein gene expression in *Pneumocystis carinii*
JOURNAL J. Infect. Dis. 171 (6), 1563-1568 (1995)
MEDLINE 95287050
REFERENCE 2 (bases 1 to 81)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1994) to the DDBJ/EMBL/GenBank databases.
Yoshikazu Nakamura, Institute of Medical Science, University of
Tokyo, Tumor Biology; Shirokanedai 4-6-1, Minato-ku, Tokyo 108,
Japan (E-mail:nak@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5307,
Fax:03-5449-5415)
FEATURES Location/Qualifiers
source 1..81
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"
gene 1..81
/gene="MSG"
5'UTR 1..81
/gene="MSG"
/note="major surface glycoprotein"
BASE COUNT 30 a 10 c 13 g 28 t
ORIGIN
1 attcatgaag taactatcaa gcttaataaa ggacccaata taggagtata ttgttaatat
61 ttatgctata tattgggatc c
//

LOCUS D31914 496 bp mRNA PL 22-SEP-1997
DEFINITION *Pneumocystis carinii* MSG mRNA for major surface glycoprotein,
5'UTR.
ACCESSION D31914
NID g2443378
KEYWORDS MSG.
SOURCE *Pneumocystis carinii* cDNA to mRNA.
ORGANISM *Pneumocystis carinii*
Eukaryotae; Fungi; Ascomycota; Archaeascomycetes;
Pneumocystidaceae; Pneumocystis.
REFERENCE 1 (sites)
AUTHORS Wada,M., Sunkin,S.M., Stringer,J.R. and Nakamura,Y.
TITLE Antigenic variation by positional control of major surface
glycoprotein gene expression in *Pneumocystis carinii*
JOURNAL J. Infect. Dis. 171 (6), 1563-1568 (1995)
MEDLINE 95287050
REFERENCE 2 (bases 1 to 496)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1994) to the DDBJ/EMBL/GenBank databases.
Yoshikazu Nakamura, Institute of Medical Science, University of
Tokyo, Tumor Biology; Shirokanedai 4-6-1, Minato-ku, Tokyo 108,
Japan (E-mail:nak@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5307,
Fax:03-5449-5415)
FEATURES Location/Qualifiers
source 1..496
/organism="Pneumocystis carinii"
/db_xref="taxon:4754"
gene 1..496
/gene="MSG"
5'UTR 1..496
/gene="MSG"
/note="major surface glycoprotein"
BASE COUNT 214 a 50 c 89 g 143 t
ORIGIN
1 caactatcaa gcttaataaa ggacccaata taggagtata ttgttaatat ttatgtata
61 tattgggatc cagaaataat atgaaatgct ttttccaaa attaataaaaa ttatagttct
121 aagaatatta atttaatttg ttccctactta tgccaatcaa agcacacatt taaaaggccag
181 agattctcaa gtaagcgcag caaatactcc gtttaaggat attgttgtta aggatgaaga
241 aatccttgct tatattttga gagaagatta taaagataaa tgcgaactaa agcttgagga
301 atattgtaaa gaattgaggg atatacatca aggattaaat aaagttcata ctatagttaa
361 agaaaattgt gatgatgaaa aacgagacgg aaaatgc当地 gaactgaaag acaaagatcaa
421 aaaggaatttga aagaggaact tgaaaaagca ttgaaaagaca taaaagatga
481 aaatttgtgaa aaatat
//